



Object-based semi-automatic approach for forest structure characterization using lidar data in heterogeneous *Pinus sylvestris* stands

C. Pascual^{a,*}, A. García-Abril^b, L.G. García-Montero^a, S. Martín-Fernández^c, W.B. Cohen^d

^a Department of Forest Engineering, Technical University of Madrid (UPM), E.T.S. I. Montes, Ciudad Universitaria s/n, 28040 Madrid, Spain

^b Department of Projects and Rural Planning, Technical University of Madrid (UPM), E.T.S. I. Montes, Ciudad Universitaria s/n, 28040 Madrid, Spain

^c Department of Economy and Forest Management, Technical University of Madrid (UPM), E.T.S. I. Montes, Ciudad Universitaria s/n, 28040 Madrid, Spain

^d Forestry Sciences Laboratory, Pacific Northwest Research Station, USDA Forest Service, 3200 SW Jefferson Way, Corvallis, OR 97311, USA

ARTICLE INFO

Article history:

Received 5 June 2007

Received in revised form 25 February 2008

Accepted 28 February 2008

Keywords:

Lidar

Forest structure

Pinus sylvestris

Mean height

Forest management

ABSTRACT

In this paper, we present a two-stage approach for characterizing the structure of *Pinus sylvestris* L. stands in forests of central Spain. The first stage was to delimit forest stands using eCognition and a digital canopy height model (DCHM) derived from lidar data. The polygons were then clustered (*k*-means algorithm) into forest structure types based on the DCHM data within forest stands. Hypsographs of each polygon and field data validated the separability of structure types. In the study area, 112 polygons of *Pinus sylvestris* were segmented and classified into five forest structure types, ranging from high dense forest canopy (850 trees ha⁻¹ and Lorey's height of 17.4 m) to scarce tree coverage (60 tree ha⁻¹ and Lorey's height of 9.7 m). Our results indicate that the best variables for the definition and characterization of forest structure in these forests are the median and standard deviation (S.D.), both derived from lidar data. In these forest types, lidar median height and standard deviation (S.D.) varied from 15.8 m (S.D. of 5.6 m) to 2.6 m (S.D. of 4.5 m). The present approach could have an operational application in the inventory procedure and forest management plans.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Forest structure can be defined by size, age, and species distributions of living and dead vegetation, often with a focus on the tree component (Spies and Franklin, 1991; Poage and Tappeiner, 2005). Structure includes both vertical (e.g. number of tree layers, understory vegetation) and horizontal features (e.g. spatial pattern of trees, gaps) as well as species richness (Maltamo et al., 2005). The conventional organisational level for forest structure is the stand, which in unmanaged systems represent the synthesis of ecological and environmental factors. Smith et al. (1997) define the stand as a contiguous group of trees sufficiently uniform in species composition, arrangement of age classes, site quality, and condition, to be a distinguishable unit. In many settings, forest stands are additionally defined by forestry activities as operational units in forest planning and management (Nyland, 1996; Holmstrom, 2002; Leckie et al., 2003; Tiede et al., 2004; Maltamo et al., 2005).

For example, stands are the basic survey units of most Spanish forest management plans. In Spain, forest inventories follow a traditional procedure. They are based on a systematic sampling

design of field plots with areas ranging from 300 to 1000 m², in which basic tree attributes dbh (diameter at breast height), height, crown size, etc. are measured. Next, stand-level mean values of biophysical variables such as dominant height, basal area, stem number, volume and growth are calculated as average field plot measurements. Finally, these stands are grouped in broader forest structure types, based on stand-level characteristics, to be treated as units for the purposes of management and forestry applications.

For management purposes, stand boundaries have traditionally been delineated on aerial photographs by means of human pattern recognition, and then interpreted photogrammetrically with the support of local field knowledge and observation (Franklin, 2001). Although this traditional approach is quite useful, it is also time consuming, and the stand delineation process is highly subjective (Skidmore, 1989; Franklin, 2001). However, the identification and spatial delineation of clusters of similar trees in the forest stand inventory improves the precision of stand-level growth and yield predictions and stand-level inventories (Magnussen et al., 2006).

Consequently, remote sensing, via image segmentation and statistical modelling, has been advanced to assist in forest surveys. Leckie et al. (2003) suggest that semi-automated, computer-assisted interpretation of digital imagery offers a possible method of acquiring some or all of the desired information, reducing time and costs, and increasing consistency. This is accomplished with

* Corresponding author. Tel.: +34 91 3367120; fax: +34 91 5439557.

E-mail address: c.pascual@upm.es (C. Pascual).

the use of different digital imagery and, to a lesser extent, with lidar (light detection and ranging) data (Diedershausen et al., 2004; Tiede et al., 2004; Magnussen et al., 2006).

Most forestry lidar studies have concentrated on plot-level or stand-level estimation of such attributes as mean height, volume or basal area, canopy closure and others (Jensen et al., 2006; van Aardt et al., 2006; Wallerman and Holmgren, 2007) with a focus on reducing field work effort and increasing accuracy. Naesset (2004) estimated principal stand characteristics with higher precision using lidar data than applying conventional methods in forest inventory in areas with terrain slopes of up to 30°–35°, and no mixed forest canopy formations. Naesset et al. (2004) indicate that area-based approaches to estimate forest stand variables from laser scanner data have matured and are now implemented in operational projects in Scandinavian countries.

Lidar data have also highlighted the relevance of the canopy height as an attribute for characterising forest structure. Variation in tree height (e.g., standard deviation of height) is primarily focused on characterizing the vertical structure of a forest canopy (Zenner, 2000; McElhinny et al., 2005). Lefsky et al. (1999), Means et al. (1999) Harding et al. (2001) and Parker and Russ (2004) developed canopy height profiles from lidar data to synthesize the three-dimensional distribution of forest canopies. Zimble et al. (2003) used lidar-derived tree height variances to distinguish between single-story and multi-story classes. Parker and Russ (2004) used the standard deviation of lidar-derived height to describe the topography of the outer canopy.

According to Lefsky et al. (2005), mean height and height variability derived from lidar data are strongly related to canopy indices related to stand structure. These authors work in plots with young, mature and old-growth formations, in largely closed canopy stands which by definition are not heterogeneous in terms

of percentage of forest cover. However, in horizontally heterogeneous forests (i.e., forests with variable amounts of tree cover), the use of mean height poses problems for the classification of different forest stands (Pascual, 2006; Pascual et al., 2006). These authors therefore suggest analyzing statistics other than mean height to obtain a better classification of these heterogeneous forest stands.

In this context, this study focuses on lidar data to study forest structure in order to be of practical application in forest management. The general objective was to characterize the forest structure based on lidar height distributions, supplemented by field data to support the lidar interpretations. This involved three secondary objectives. The first was to define forest stand boundaries and cluster these stands into forest structure types based on the lidar height distributions. The second was to select a few lidar height summaries to make the present approach more effective. The third was to propose a methodological approach that would ultimately reduce the need for expensive fieldwork and which would be easy to implement in operational forest management.

We therefore decided to work with the lidar derived digital canopy height model (DCHM) directly provided by commercial firms, in order to establish whether it was possible to use the lidar processed information to implement a feasible methodological approach in areas that fulfil the requirements proposed by Naesset (2004).

2. Materials and methods

2.1. Study site

A 127.10 ha (1293 m × 983 m) area on the western slopes of the Fuenfría Valley (40°45'N, 4°5'W) in central Spain was selected as

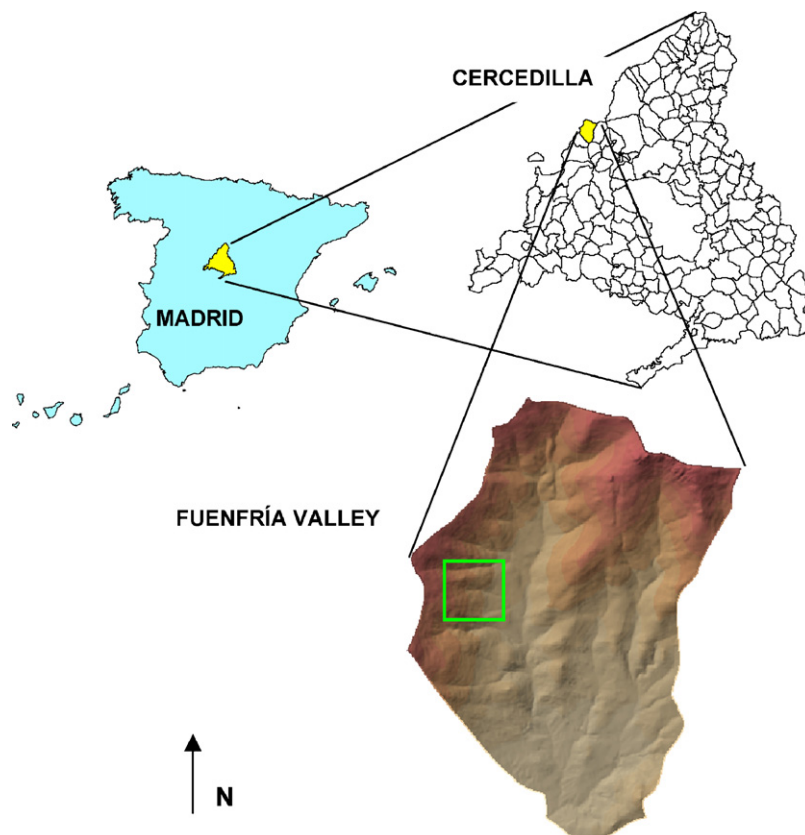


Fig. 1. Study site. Fuenfría Valley in the village of Cercedilla, northwest of Madrid (Spain).

the study area. The Fuenfría Valley is located in the northwest portion of the Madrid region (Fig. 1). The predominant forest is Scots pine (*Pinus sylvestris* L.) with abundant shrubs (*Cytisus scoparius* (L.) Link., *C. oromediterraneus* Rivas Mart. et al., *Genista florida* L.) in some areas. There are small pastures on the lower slopes of the hillside. In the north sector of the study site there is an extensive rocky area. The site has a mean annual temperature of 9.4 °C and precipitation averages 1180 mm/year. Elevations range between 1310 and 1790 m above sea level, with slopes of between 20% and 45%. The general aspect of the study site is east.

2.2. Lidar data and pre-processing

A small-footprint lidar dataset was acquired by Toposys GmbH over the study area in August, 2002. The Toposys II lidar system recorded first and last returns with a footprint diameter of 0.95 m. Average point density was 5 points m^{-2} . The raw data (x , y , and z coordinates) was processed into two digital elevation models by TopoSys using as interpolation algorithm a special local adaptive median filter developed by the data provider. The digital surface model (DSM) was processed using the first pulse reflections, and the digital terrain model (DTM) was constructed using the last returns. Filtering algorithms were used to identify canopy and ground surface returns for an output pixel resolution of 1 m horizontal and 0.1 m vertical resolution. According to Toposys calculations, the DSM and DTM, horizontal positional accuracy was 0.5 m and vertical accuracy was 0.15 m.

To obtain a DCHM, the DTM was subtracted from the DSM. Both the DTM and DCHM were validated before use by means of land surveying with total station in 19 points and ground-based height measurements of 102 trees. The vertical accuracies, (Root Mean Square Error, R.M.S.E.s) obtained for the DTM in open areas and for the DCHM under forest canopy were 0.30 and 1.3 m, respectively. These accuracies were acceptable for this study, and were in agreement with previous studies. For example, Clark et al. (2004) reported R.M.S.E.s for DTMs ranging from 0.06 to 0.61 m, and for DCHMs ranging from 0.23 to 2.41 m in tropical landscapes.

2.3. Mapping forest structure

The process of mapping forest structure for this study involved several steps. These include delineating and classifying polygons and defining the classified polygons with DCHM hypsographs and additional field data.

2.3.1. Polygon delineation and classification

The aim of this step was to use an object-oriented segmentation approach to delineate forest stand boundaries (polygons) and classify them in the DCHM, based on lidar height distributions.

Polygons were delineated from the lidar-derived DCHM using eCognition 4.0 software (Definiens Imaging GmbH, Munich 2004). This is an object-based image analysis package which applies a spatial clustering technique (Haralick and Shapiro, 1985). This algorithm identifies geographical features using the scale and homogeneity parameters, which were obtained in this study from the lidar DCHM. According to Suarez et al. (2005) scale relates to the minimum size required to identify a particular object, which depends on the resolution of the images. Homogeneity is described by a mutually exclusive interaction between colour and shape. Colour refers to the spectral response of the objects (lidar height in this study), whereas shape is divided into two equally exclusive properties: smoothness and compactness which respectively define the boundaries of the polygons (objects) and their transition to others.

Three consecutive segmentations were applied to the lidar DCHM. A first segmentation with a scale parameter of 30 was

derived, with 0.5 and 0.3 as the shape and smoothness parameters, respectively. These objects were later aggregated into the levels of higher hierarchical scales 50 and 70 by using the same smoothness parameter (0.3), with 0.4 as the shape values for the second and third segmentation. These values were demonstrated as the most feasible for the assessment of objects representing forest stands with a reasonably homogeneous structural typology.

A total of 146 polygons were segmented, 112 containing Scots pines, and 34 consisting of non-forest components such as pasture, shrubs, rocks and bare soil. The 112 forest polygons were grouped into five structure types by a *k-means* cluster analysis. Separate cluster analyses were performed on two different combinations of variables derived from summaries of the fundamental DCHM data within each polygon: (1) mean and standard deviation (S.D.) of height (CombV1), and (2) median and S.D. of height (CombV2). In previous work (García-Abril et al., 2006; Pascual, 2006), we have studied several indices and variables derived from the lidar-DCHM: relative gap surface, landscape ecology metrics (i.e. Shannon index, fractal dimension, contagion index) voxels and texture, but these were rejected as they did not contribute to cluster discrimination. The coefficient of variation (CV) (i.e. standard deviation divided by the mean) was not considered as an entry variable as it depends on the rest. According to Hair et al. (1995) there is no objective procedure for establishing the number of clusters. These authors suggest obtaining various cluster solutions and deciding based on *a priori* criteria, experience or theoretical foundations. In this study, the decision to base the analysis on five structure types was an iterative process aided by the expert opinion of forest management personnel in the area.

Individual polygons were assigned to the different clusters using the sequential threshold method, where distances in cluster seeds were sorted, and observations of the distances between them taken at constant intervals. Analysis of variance was used to test the statistical significance of the forest structure types derived from the cluster analysis. Euclidean distances between cluster centroids were also used as an indicator of the proximity of cluster groupings (Hair et al., 1995; Levina, 2003).

2.3.2. Cluster hypsographs and field-based description

The two sets of polygon assignment to clusters were evaluated for their relative value in distinguishing among the five forest structure types. As no tests exist to measure the weight of the variables used to define the clusters in the cluster statistics technique (Hair et al., 1995), we decided to compare the classifications using hypsographs. Hypsographs are the cumulative distribution of canopy heights as a function of proportional area within each polygon. Such a graph has been used in the description of landform surfaces and their development (Strahler, 1952). Hypsographs are related to canopy height profiles, and are both powerful tools for synthesizing the three-dimensional distribution of forest canopies, and have been widely used to analyze stand structure, either derived from lidar data (Lefsky et al., 1999; Harding et al., 2001; Maltamo et al., 2005) or from the DCHM (Parker and Russ, 2004). Hypsographs of each forest polygon were derived from distributions of heights contained within the lidar DCHM.

Further analysis involved summarizing the hypsographs into percentiles; i.e., heights at which 10%, 25%, 50%, 75% and 90% of the polygon surface area occurs within each polygon ($H_{10\%}$, $H_{25\%}$, $H_{50\%}$, $H_{75\%}$ and $H_{90\%}$, respectively). ANOVA and Kruskal–Wallis tests with Tukey's method for *post hoc* analysis were used to test whether indices describing canopy height distributions ($H_{10\%}$, $H_{25\%}$, $H_{50\%}$, $H_{75\%}$ and $H_{90\%}$) varied significantly among forest structure types for both variable combinations (CombV1 or CombV2). These results were useful for determining which variable combination provided better separability of structure

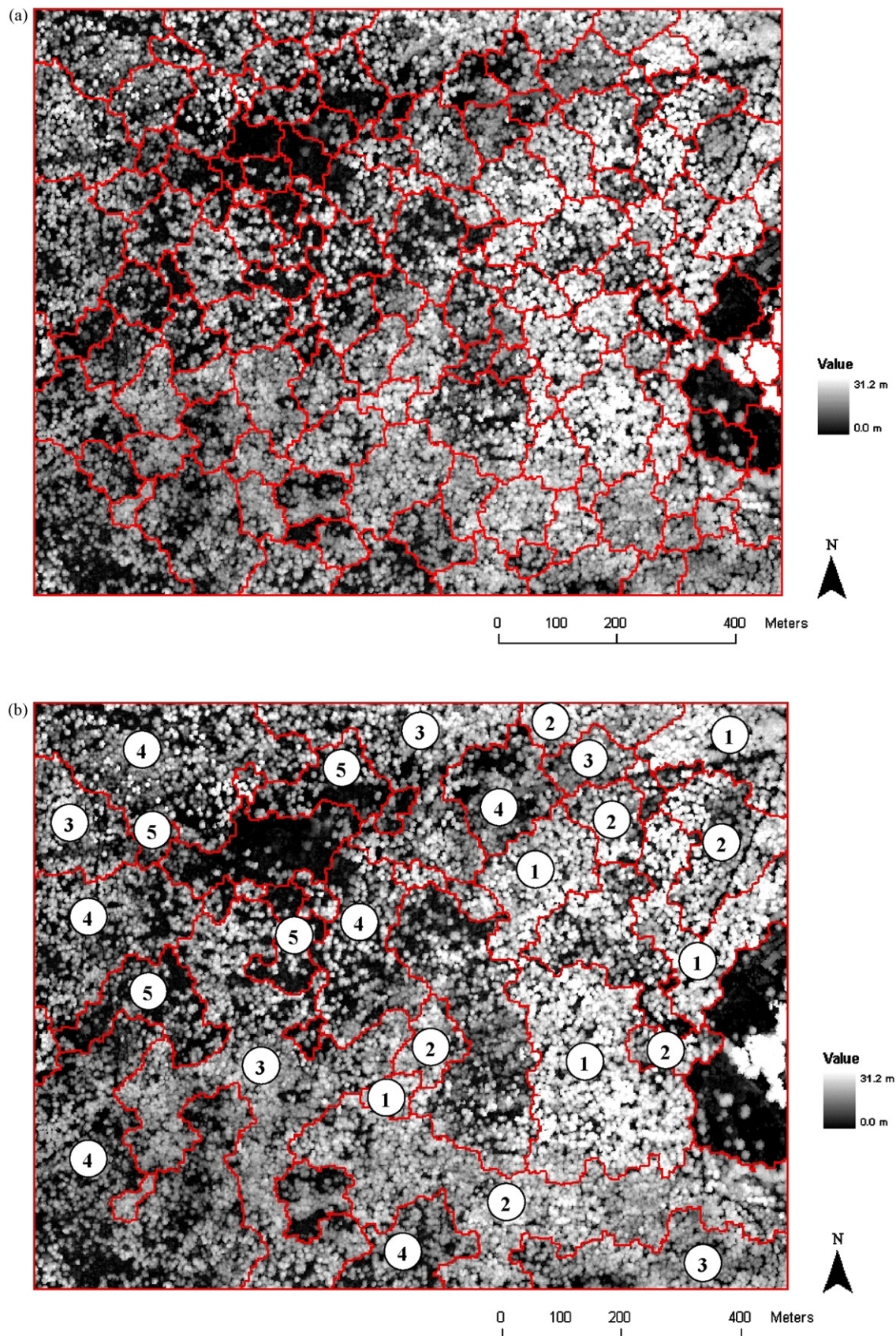


Fig. 2. (a) The 112 object-based segmented polygons using eCognition. (b) Results of the cluster analysis (*k*-means). Numbers inside the polygons indicate the forest structure type (1, 2, 3, 4 or 5) to which each polygon has been ascribed by cluster analysis. Polygons without numbers correspond to non-forest stands.

Table 1

Descriptive statistics (mean and standard deviation) of entry variables in the two cluster combinations (CombV1 and CombV2)

	Lidar height variables	Forest type 1	Forest type 2	Forest type 3	Forest type 4	Forest type 5
Comb V1	Mean height	14.3 (1.1)	11.4 (0.8)	9.1 (0.7)	6.7 (0.7)	3.8 (0.9)
	S.D. ¹ height	5.6 (0.9)	4.6 (0.9)	5.5 (0.7)	5.1 (0.7)	4.1 (0.7)
	No. of members	27	26	28	23	8
Comb V2	Median height	15.8 (0.8)	13.0 (0.8)	10.3 (0.7)	7.3 (0.9)	2.6 (1.4)
	S.D. ¹ height	5.6 (1.0)	4.9 (1.0)	5.0 (1.0)	5.1 (0.7)	4.5 (1.1)
	No. of members	22	23	29	25	13

S.D.¹ = standard deviation. Parenthetical values are standard deviations of mean, median and S.D.¹ of height for polygons in each forest type.**Table 2**

Euclidean distance among cluster (forest types) centres for CombV1 and CombV2

	CombV1					CombV2			
	1 ⁺	2 ⁺	3 ⁺	4 ⁺		1 ⁺	2 ⁺	3 ⁺	4 ⁺
2 ⁺	2.17	–	–	–	2 ⁺	2.08	–	–	–
3 ⁺	3.72	1.80	–	–	3 ⁺	3.97	1.93	–	–
4 ⁺	5.41	3.35	1.71	–	4 ⁺	6.02	3.99	2.06	–
5 ⁺	7.55	5.44	3.88	2.17	5 ⁺	9.42	7.37	5.45	3.41

1⁺, 2⁺, 3⁺, 4⁺, 5⁺ stand for forest types 1, 2, 3, 4 and 5, respectively.

types. All statistical analyses were done with STATISTICA v. 6.1 software (StatSoft, Inc., Tulsa, OK, 2004).

We collected field data to provide an independent assessment of the forest structure types. Ten plots were established, consisting of two different sets; one with two plots of 2400 m² each and another with 8 plots of 1260 m². Two field plots were located in each forest type. They were inventoried during July 2003. The DBH (diameter at breast height) and height of all trees on each plot were measured, the latter with a Vertex III hypsometer. These data were used to describe the five structure types. Plot centres were geo-referenced using Differential Global Positioning System (GPS). A Trimble Geoexplorer 3 receiver observing the C/A-code and carrier phase (L1) was used. Collection of data lasted about 40–60 min for each plot, with a 5-s logging rate. Post-processing was computed using GPS Pathfinder 2.70 software. Correction data were downloaded from the Crustal Dynamics Data Information System (CDDIS) using a base station 30 km away. According to the positional errors reported by the software, the horizontal accuracy of the plot coordinates ranged from 0.6 to 2.8 m.

3. Results

3.1. Polygon segmentation and classification

Lidar data segmentation with eCognition provided 112 polygons (Fig. 2a) varying in size from 0.138 to 3.982 ha, with an average area of 0.926 ha (S.D. of 0.666 ha). The clustering of polygons into five forest structure types based on either mean (CombV1) or median (CombV2) and standard deviation of DCHM within-polygon heights (Table 1), revealed that both combinations were able to separate all 5 types (analysis of variance F -ratios between cluster centres: (i) CombV1 (mean $F = 376.58$; $p < 0.001$ and S.D. $F = 8.58$; $p < 0.001$), (ii) CombV2 (median $F = 526.91$;

$p < 0.001$ and S.D. $F = 3.67$; $p < 0.001$)), however, CombV2 was slightly better in doing so (Table 2).

3.2. Hypsographs

The hypsographs and related height histograms reveal meaningful differences among the five forest structure types (Fig. 3). This is confirmed by analysis of the hypsograph percentiles (Tables 3a and 3b). Both percentile indices for CombV2 and CombV1 were able to discriminate all structure types, whereas for CombV2, only the $H_{10\%}$ percentile was not statistically significant for distinguishing between structure types 1 and 2 and types 4 and 5. This result, together with Euclidean distance separations among cluster centroids for CombV2 (Table 2), suggests that median and standard deviation of the DCHM are better suited to distinguishing among the five *P. sylvestris* structure types defined in this study.

3.3. Field verification of structural differences

The data collected from the 10 plots in order to describe the five forest structure types in the study site, independent of the lidar data, reveal meaningful differences among types.

Type 1: Uneven-aged forest (multilayered canopy) with very high crown cover. These forest stands are located in the lowest part of the hillside, between 1330 and 1470 m in the study area (Fig. 2b). The slope ranges from 20% to 30% and the general aspect is northeast and southeast. This forest type corresponds to a multilayered, uneven-aged Scots pine formation (Fig. 4a). Crown cover ranges between 75% and 85%, and density is over 850 trees ha⁻¹. This forest type includes the tallest trees in the study area (Table 4 and Fig. 4a).

Type 2: Multi-diameter forest with high crown cover. These polygons are distributed between 1310 and 1600 m in the southern

Table 3aTukey HSD test significant differences ($p < 0.05$): hypsograph indices, CombV1

	Type 1	Type 2	Type 3	Type 4
Type 2	H _{25%} , H _{50%} , H _{75%} , H _{90%}			
Type 3	H _{10%} , H _{25%} , H _{50%} , H _{75%} , H _{90%}	H _{10%} , H _{25%} , H _{50%} , H _{75%}		
Type 4	H _{10%} , H _{25%} , H _{50%} , H _{75%} , H _{90%}	H _{10%} , H _{25%} , H _{50%} , H _{75%} , H _{90%}	H _{25%} , H _{50%} , H _{75%} , H _{90%}	
Type 5	H _{10%} , H _{25%} , H _{50%} , H _{75%} , H _{90%}	H _{10%} , H _{25%} , H _{50%} , H _{75%} , H _{90%}	H _{25%} , H _{50%} , H _{75%} , H _{90%}	H _{50%} , H _{75%} , H _{90%}

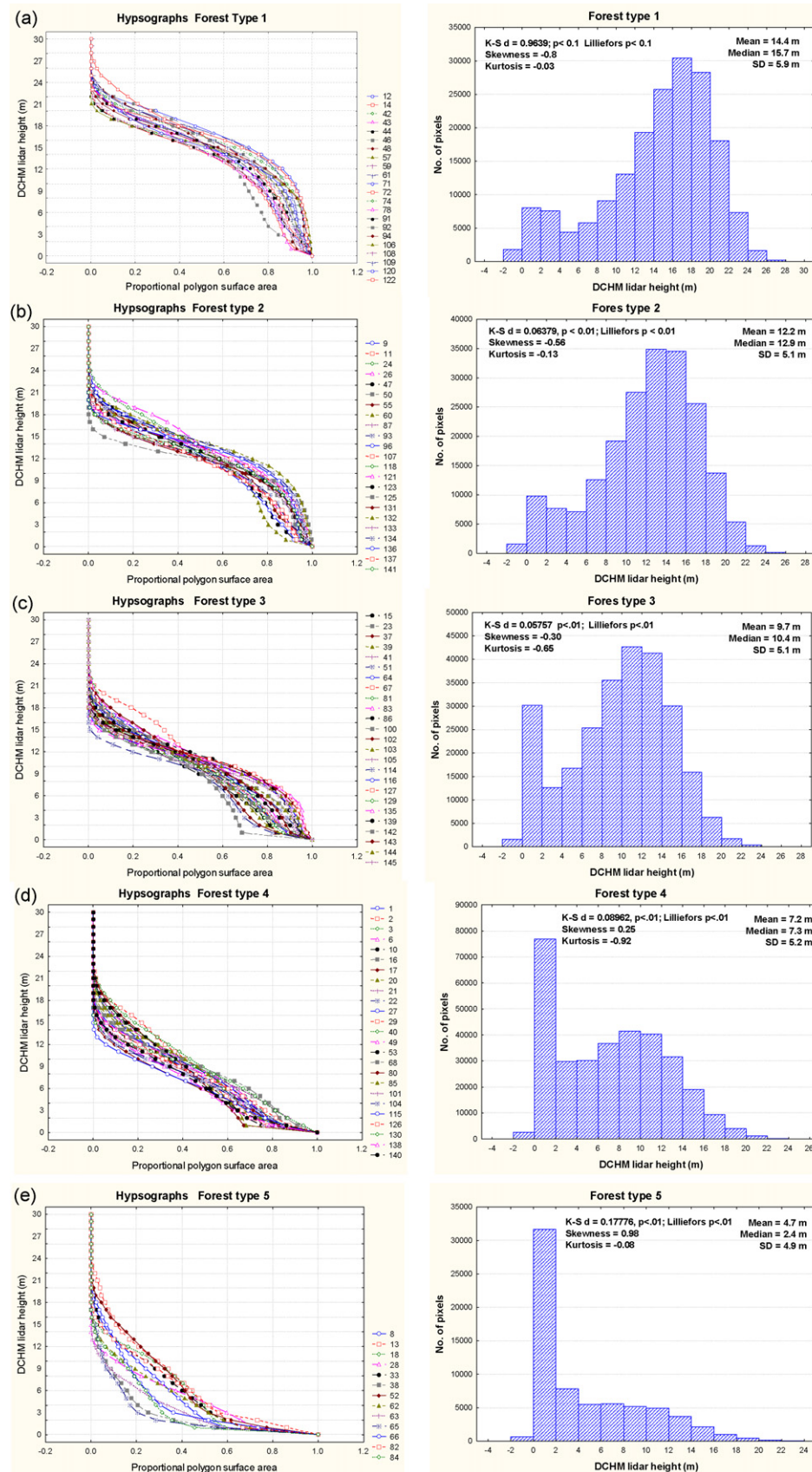


Fig. 3. This figure is based on the 112 segmented polygons. Hypsographs of DCHM data for each forest stand polygon, by forest type. Numbers represent the identification code for each polygon (left); DCHM lidar height histograms (right).

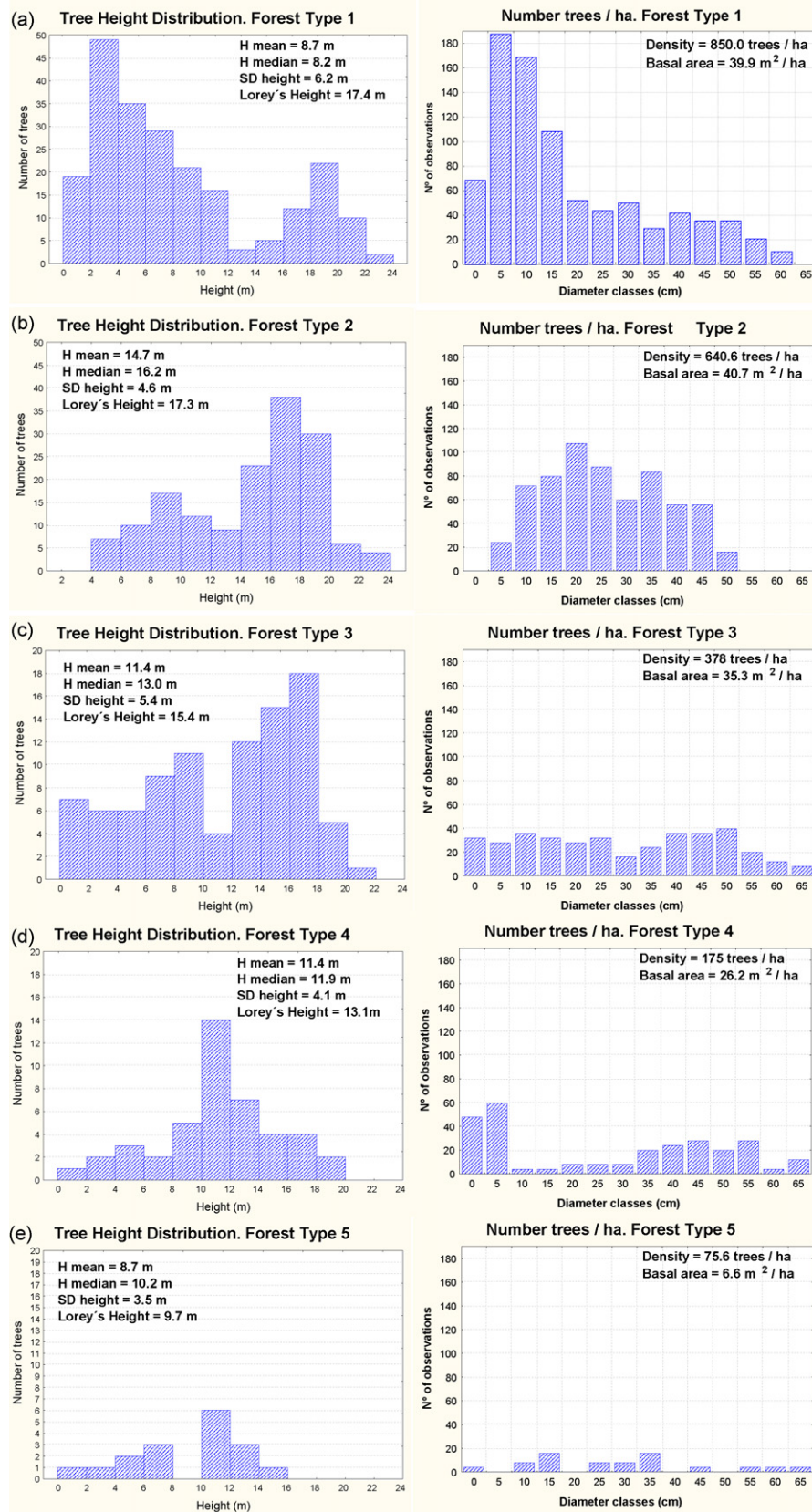


Fig. 4. Field measurements for the 10 plots used to describe forest structure types. Tree height distribution (left); diameter class distribution (right). Lorey's height is a mean tree height weighted by the normal section of trees.

Table 3bTukey HSD test significant differences ($p < 0.05$): hypsograph indices, CombV2

	Type 1	Type 2	Type 3	Type 4
Type 2	H _{25%} , H _{50%} , H _{75%} , H _{90%}			
Type 3	H _{10%} , H _{25%} , H _{50%} , H _{75%} , H _{90%}	H _{10%} , H _{25%} , H _{50%} , H _{75%} , H _{90%}		
Type 4	H _{10%} , H _{25%} , H _{50%} , H _{75%} , H _{90%}	H _{10%} , H _{25%} , H _{50%} , H _{75%} , H _{90%}	H _{10%} , H _{25%} , H _{50%} , H _{75%} , H _{90%}	
Type 5	H _{10%} , H _{25%} , H _{50%} , H _{75%} , H _{90%}	H _{10%} , H _{25%} , H _{50%} , H _{75%} , H _{90%}	H _{10%} , H _{25%} , H _{50%} , H _{75%} , H _{90%}	H _{25%} , H _{50%} , H _{75%} , H _{90%}

Table 4

Forest attributes from 10 field plot measurements in the five forest structure types

Forest type	Mean height (m)	S.D. of height (m)	Lorey's height (m)	Basal area (m ² ha ⁻¹)	Density (trees ha ⁻¹)
1	9.9	6.2	17.4	39.9	850
2	14.7	4.6	17.3	40.7	640
3	11.4	5.4	15.4	35.3	378
4	11.4	4.1	13.1	26.2	175
5	8.7	3.5	9.7	6.6	76

Lorey's height is a mean tree height weighted by the normal section of trees.

portion of the study area with some discontinuous polygons in the north sector (Fig. 2b). The general aspect is east, and slopes range between 20% and 50%, with a mean of over 35%. This forest type can be described as having a multi-diameter distribution and a two-story vertical distribution (Fig. 4b). Canopy cover is over 65%–70% and density is 640 trees ha⁻¹. Trees included in this forest type have slightly lower height and diameter than in the previous one (Table 4) (Fig. 3b).

Type 3: Multi-diameter forest with medium crown cover. This type occurs discontinuously across the elevation gradient of the study area (1310–1790 m) (Fig. 2b). Predominant aspect is east. The mean slope is 40%, crown cover is over 55% and density is 178 trees ha⁻¹. This type of forest has a multi-diameter distribution, but is less dense than type 2 above (Fig. 4c). In some polygons the pines form clumps of trees.

Type 4: Even-aged forest (single-story) with low crown cover. These stands are distributed throughout the higher elevations of the study area (1500–1790 m), which has a predominantly eastern orientation. Slopes range between 40% and 55%. The distribution of diameter classes is close to an even-aged formation and height distribution represents a single-story condition (Fig. 4d). This forest type includes mature trees of greater diameter but with a slightly lower height and larger crown diameters than other types (Fig. 4d). Crown cover is relatively low, and is generally less than 40% and density is 175 trees ha⁻¹. In this type a subtype was identified, consisting of a two-layered structure with emergent trees over a regenerating stand.

Type 5: Zones with scarce tree coverage. This type consists of dense coverage of shrubs (*Rosa* spp., *Adenocarpus hispanicus* (Lam.) DC., *Cytisus scoparius* (L.) Link, *Pteridium aquilinum* (L.) Kuhn and *G. florida* L.) under isolated pine trees. Crown cover is between 10% and 15%, with very low density (75 trees ha⁻¹) (Fig. 4e). These polygons are located at the highest elevations (1550–1750 m), with a mean slope of up to 40% and a predominantly northern or eastern aspect.

4. Discussion

Modern remote sensing tools, lidar data and object-oriented segmentation approaches are making possible to automate many of the processes involved in forest structure stand delineation. In this study we demonstrate the utility of airborne lidar data for this purpose, as lidar directly provides forest canopy height that has traditionally been estimated using a combination of aerial photo interpretation and photogrammetry.

The present study proposes a three-step methodological approach for forest structure characterization. The first step is

the segmentation of the laser scanner DCHM in forest stands; the second is to cluster these stands into forest structure types based on the lidar height summaries; and the final step is to validate the procedure with field data and hypsographs.

Some authors propose automated forest stand segmentation procedures in homogeneous forest formations (Diedershausen et al., 2004). Our results show the utility of the combination of eCognition algorithms with knowledge of the study area in order to obtain an automated segmentation of forest stands in horizontally heterogeneous forest formations.

Lidar DCHMs provide a vast amount of information on forest structure stands that can be used in a variety of ways to cluster the stands into forest structure types. Pascual (2006) and Garcia-Abril et al. (2006) carried out several cluster trials considering multiple combinations of entry variables derived from a binned and non-binned lidar DCHM: mean height, median, S.D., relative gap surface area (i.e. number of grid of zero height divided by total number of polygon grids), mean and S.D. of texture in a 5 × 5 window, and different landscape indices (fractal dimension, Shannon index, contagion index and others). These authors have proved that most of these variables were redundant in the study area.

Our results showed that the mean and standard deviation of height provided valuable characterization of the forest structure for types 1 and 2, which have height distributions which tend to be closer to normality (Fig. 3a and b). This result is in accordance with (Zimble et al., 2003; Lefsky et al., 2005). However the median and standard deviation were better at distinguishing structure types 3, 4 and 5, which had more pronounced non-normal height distributions (Fig. 3c–e). This is because, for such distributions, the most representative central value is the median (Quinn and Keough, 2002). Of course, for normal distributions, mean and median are nearly identical. We therefore conclude that, in general, median height derived from DCHMs is better related to structure in forests with heterogeneous height distributions.

In conclusion, our results show that median and S.D. of height derived from lidar DCHM were useful for distinguishing among horizontally heterogeneous forest structure types based on cluster analysis. They met the requirements for this present approach, in that they were flexible and easy to implement in operational management plans for Spanish forests. However, the original lidar point cloud might contain additional information of special interest in horizontally heterogeneous forests that could be analyzed in future research.

Finally, in order to apply the results of the present methodological approach in forest inventories, we considered it especially useful to validate this approach to creating a model of forest structure types with field data measurements. Our proposal for

validation based on hypsographs and percentiles allow a quick validation.

Acknowledgments

Our thanks to M.A. Grande, J.A. Manzanera, J. Velazquez, E. Blanco and E. Jiménez for their collaboration in field data measurements; to P. Díaz, E. Ayuga, J.L. Hernanz, F. Marcos and I. Izquierdo for their support and collaboration and to L. Arroyo for her teaching of eCognition. Financial support for the study was provided by the Madrid Education and Culture Council through Project GR/AMB/0267/2004 and by the Technical University of Madrid (UPM) through Project N° CCG06-UPM/9999-440. We wish to acknowledge the Rey Juan Carlos University for financing 3 months of study for C. Pascual as a visiting scholar at the Laboratory of Applications of Remote Sensing in Ecology (LARSE), in Corvallis (OR), USA.

Appendix A. Abbreviations

DSM	Digital surface model
DTM	Digital terrain model
DCHM	Digital canopy height model
CombV1	Mean and standard deviation of lidar height as combination of entry variables for cluster analysis
CombV2	Median and standard deviation of lidar height as combination of entry variables for cluster analysis
S.D.	Standard deviation
CV	Coefficient of variation
R.M.S.E.	Root mean square error

References

- Clark, M.L., Clark, D.B., Roberts, D.A., 2004. Small-footprint lidar estimation of sub-canopy elevation and tree height in a tropical rain forest landscape. *Rem. Sens. Environ.* 91, 68–89.
- Diedershausen, O., Koch, B., Weinacker, H., 2004. Automatic segmentation and characterization of forest stand parameters using airborne lidar data, multispectral and fogis data. In: Thies, M., Koch, B., Spiecker, H., Weinacker, H. (Eds.), *Laser-Scanners for Forest and Landscape Assessment*. Institute for Forest Growth Institute for Remote Sensing and Landscape Information Systems. Albert Ludwigs University Freiburg, Germany, pp. 208–212.
- Franklin, S.E., 2001. Remote sensing for sustainable forest management. Lewis Publisher, Boca Raton, USA.
- García-Abril, A., Pascual, C., Ruiz, M.A., Velarde, M.D., Blanco, E., Velázquez, J., Cocero, D., 2006. Caracterización de los tipos de estructura forestal mediante índices de ecología del paisaje aplicados sobre datos lidar. In: XII Congreso Nacional de Tecnologías de la Información, Granada.
- Hair, J.F., Anderson, R.E., Tatham, R.L., Black, W.C., 1995. *Análisis Multivariante*. Prentice-Hall International, London.
- Haralick, R., Shapiro, L., 1985. Image segmentation techniques. *Comp. Vis. Graph. Image Proc.* 29, 100–132.
- Harding, D.J., Lefsky, M.A., Parker, G.G., Blair, J.B., 2001. Laser altimeter canopy height profiles: methods and validation for closed-canopy, broadleaf forests. *Rem. Sens. Environ.* 76, 283–297.
- Holmstrom, H., 2002. Estimation of single-tree characteristics using the kNN method and plotwise aerial photograph interpretations. *Forest Ecol. Manage.* 167, 303–314.
- Jensen, J.L.R., Humes, K.S., Conner, T., Williams, C.J., DeGroot, J., 2006. Estimation of biophysical characteristics for highly variable mixed-conifer stands using small-footprint lidar. *Can. J. For. Res.* 36, 1129–1138.
- Leckie, D.G., Gougeon, F.A., Walsworth, N., Paradine, D., 2003. Stand delineation and composition estimation using semi-automated individual tree crown analysis. *Rem. Sens. Environ.* 85, 355–369.
- Lefsky, M.A., Cohen, W.B., Acker, S.A., Parker, G.G., Spies, T.A., Harding, D., 1999. Lidar Remote Sensing of the Canopy Structure and Biophysical Properties of Douglas-Fir Western Hemlock Forests. *Rem. Sens. Environ.* 70, 339–361.
- Lefsky, M.A., Hudak, A.T., Cohen, W.B., Acker, S.A., 2005. Patterns of covariance between forest stand and canopy structure in the Pacific Northwest. *Rem. Sens. Environ.* 95, 517–531.
- Levia, D.F., 2003. Winter stemflow leaching of nutrient-ions from deciduous canopy trees in relation to meteorological conditions. *Agric. For. Meteorol.* 117, 39–51.
- Magnussen, S., Allard, D., Wulder, M.A., 2006. Poisson Voronoi tiling for finding clusters in spatial point patterns. *Scand. J. For. Res.* 21, 239–248.
- Maltamo, M., Packalen, P., Yu, X., Eerikainen, K., Hyyppä, J., Pitkanen, J., 2005. Identifying and quantifying structural characteristics of heterogeneous boreal forests using laser scanner data. *Forest Ecol. Manage.* 216, 41–50.
- McElhinny, C., Gibbons, P., Brack, C., Bauhus, J., 2005. Forest and woodland stand structural complexity: its definition and measurement. *Forest Ecol. Manage.* 218, 1–24.
- Means, J.E., Acker, S.A., Harding, D.J., Blair, J.B., Lefsky, M.A., Cohen, W.B., Harmon, M.E., McKee, W.A., 1999. Use of large-footprint scanning airborne lidar to estimate forest stand characteristics in the western cascades of Oregon. *Rem. Sens. Environ.* 67, 298–308.
- Naesset, E., 2004. Practical large-scale forest stand inventory using a small-footprint airborne scanning laser. *Scand. J. For. Res.* 19, 164–179.
- Naesset, E., Gobakken, T., Holmgren, J., Hyyppä, H., Hyyppä, J., Maltamo, M., Nilsson, M., Olsson, H., Persson, A., Soderman, U., 2004. Laser scanning of forest resources: the Nordic experience. *Scand. J. For. Res.* 19, 482–499.
- Nyland, R.D., 1996. *Silviculture: Concepts and Applications*. McGraw-Hill, New York.
- Parker, G.G., Russ, M.E., 2004. The canopy surface and stand development: assessing forest canopy structure and complexity with near-surface altimetry. *Forest Ecol. Manage.* 189, 307–315.
- Pascual, C., 2006. *Análisis de la estructura forestal mediante teledetección: LiDAR (Light Detection And Ranging) e imágenes de satélite*. Ph.D. thesis. Technical University of Madrid (UPM), Madrid.
- Pascual, C., García-Abril, A., Martín-Fernández, S., García-Montero, L.G., 2006. Application of LIDAR data to characterize structure types of multifunctional *Pinus sylvestris* L. stands. In: Reynolds, K.M. (Ed.), *Sustainable Forestry in Theory and Practice: Recent Advances in Inventory and Monitoring, Statistics and Modeling, Information and Knowledge Management, and Policy Science*. USDA For. Serv. Gen. Tech. Rep. PNW-GTR-688. Pacific Northwest Research Station, Portland (CD-ROM).
- Poage, N.J., Tappeiner, J.C., 2005. Tree species and size structure of old-growth Douglas-fir forests in central western Oregon. USA. *Forest Ecol. Manage.* 204, 329–343.
- Quinn, G.P., Keough, M.J., 2002. *Experimental Design and Data Analysis for Biologist*. Cambridge University Press, Cambridge.
- Skidmore, A.K., 1989. An expert system classifies eucalypt forest types using Thematic Mapper data and a digital terrain model. *Photogram. Eng. Rem. Sens.* 55, 1449–1464.
- Smith, D.M., Larson, B.C., Kelty, M.J., Ashton, P.M.S., 1997. *Applied Forest Ecology*. John Wiley & Sons, New York.
- Spies, T.A., Franklin, J.F., 1991. The structure of natural young, mature, and old-growth Douglas-fir forests in Oregon and Washington. In: Ruggerio, L.F., Aubry, K.B., Carey, A.B., Huff, M.H. (Eds.), *Wildlife and Vegetation of Unmanaged Douglas-fir Forests*. USDA For. Serv. Gen. Tech. Rep. PNW-285. Pacific Northwest Research Station, Portland, pp. 90–109.
- Strahler, A.N., 1952. Hypsometric (area-altitude) analysis of erosional topography. *Bull. Geol. Soc. Am.* 63, 1117–1142.
- Suarez, J., Ontiveros, C., Smith, S., Snape, S., 2005. Use of airborne LiDAR and aerial photography in the estimation of individual tree heights in forestry. *Comp. Geosci.* 31, 253–262.
- Tiede, D., Blaschke, T., Heurich, M., 2004. Object-based semi automatic mapping of forest stands with laser scanner and multi-spectral data. In: Thies, M., Koch, B., Spiecker, H., Weinacker, H. (Eds.), *Laser-Scanners for Forest and Landscape Assessment*. Institute for Forest Growth Institute for Remote Sensing and Landscape Information Systems. Albert Ludwigs University Freiburg, pp. 328–333.
- van Aardt, J.A.N., Wynne, R.H., Oderwald, R.G., 2006. Forest volume and biomass estimation using small-footprint lidar-distributional parameters on a per-segment basis. *For. Sci.* 52, 636–649.
- Wallerman, J., Holmgren, J., 2007. Estimating field-plot data of forest stands using airborne laser scanning and SPOT HRG data. *Rem. Sens. Environ.* 100, 501–508.
- Zenner, E.K., 2000. Do residual trees increase structural complexity in Pacific northwest coniferous forests? *Ecol. Appl.* 10, 800–810.
- Zimble, D.A., Evans, D.L., Carlson, G.C., Parker, R.C., Grado, S.C., Gerard, P.D., 2003. Characterizing vertical forest structure using small-footprint airborne LiDAR. *Rem. Sens. Environ.* 87, 171–182.